DOCUMENT RESUME

ED 373 301 CG 025 718

AUTHOR Beyer, Sylvia
TITLE Effects of Gender and Depression on Self-Evaluations of Performance on Academic Tasks.
PUB DATE Jul 94
NOTE 11p.; Paper presented at the Meeting of the American Psychological Society (Washington, DC, July 1994).
PUB TYPE Speeches/Conference Papers (150) -- Reports - Research/Technical (143) -- Information Analyses (070)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Attitudes; College Students; Congruence (Psychology); *Depression (Psychology); *Expectation; Higher Education; Performance; *Prediction; *Self Efficacy; *Self Evaluation (Individuals); *Sex Differences
IDENTIFIERS Beck Depression Inventory; Gender Issues

ABSTRACT

Studies have shown that women have lower expectations for success than men. This paper presents the results of four hypotheses which address depression and gender differences in self-perception: (1) depressed subjects are more accurate in their self-perceptions than non-depressed subjects; (2) self-consistency tendencies can predict gender differences in self-evaluation--women should show stronger self-consistency tendencies than men; (3) on a masculine task, women remember more of the questions they answered incorrectly than do men; and (4) subjects will be more accurate when constantly monitoring their performance than when only evaluating their overall performance. For the study, 293 females and 174 males at the University of Wisconsin-Parkside filled out the Beck Depression Inventory (BDI) and then were presented with either a feminine, masculine, or neutral gender-typed task. Subjects had to state how well they expected to do and estimate the number of correctly answered questions. Results supported the second and third hypotheses. Significantly, women more frequently revealed low confidence when they answered a question correctly than did men. Women more frequently than men remembered the questions they answered wrong, rather than correctly answered queries, and this bias is likely to adversely affect self-evaluations. Contains 66 references. (RJM)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED 373 301

CG

Effects of Gender and Depression on Self-evaluations of Performance on Academic Tasks

Sylvia Beyer

University of Wisconsin-Parkside

Paper presented at the meeting of the American Psychological Society, Washington, D. C.

July, 1994

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

S. Beyer

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☐ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

BEST COPY AVAILABLE

CG025118

Abstract

Prior research has established that gender differences in self-perceptions exist. For example, women's post-task self-evaluations of performance are lower than men's, especially on masculine gender-typed tasks (Beyer, 1990, 1991). It was hypothesized that self-consistency tendencies can partially explain gender differences in self-evaluations. According to self-consistency theory, subjects' expectancies should affect their post-task self-evaluations. The results confirmed this hypothesis. It was also assessed whether biased recall of one's performance on individual questions can partially explain gender differences in the accuracy of self-evaluations. It was found that indeed males were relatively more likely than females to recall those questions which they wrongly believed they had answered correctly.

Introduction

Women have lower expectancies of success than men (e.g., Beyer, 1987, 1990, 1991, 1992; Daubman, Heatherington, & Ahn, 1992; Elliot & Harackiewicz, 1994; Mura, 1987). Carr, Thomas, and Mednick (1985) suggested that these low expectancies are indicative of women's tendency to underestimate their abilities. Research on causal attributions has also produced evidence for women's putative underestimation of abilities. Women tend to attribute success more externally (Meehan & Overton, 1986), or more to effort rather than ability, than do men (LaNoue & Curtis, 1985; Parsons, Meece, Adler, & Kaczala, 1982). Some researchers suggest that by making external attributions for success, women are not taking credit for their performance, thereby showing a "self-derogatory" bias (Erkut, 1983). Females have also been found to have lower self-evaluations of performance than men, despite equal performance (Rustemeyer, 1982).

However, research on expectancies and attributions has not yet investigated whether women's self-perceptions are inaccurately low. For example, even though women's expectancies are lower than men's, their expectancies may be realistic, whereas men's expectancies may be overly optimistic. Thus, research has established that gender differences in expectancies and causal attributions exist but has failed to investigate the accuracy of women's and men's self-perceptions. One reason for this state of affairs is that "the study of accuracy in self-perception has been impeded by the 'criterion problem': the lack of objective criteria against which self-perceptions can be compared" (John & Robins, 1994, p. 206).

Beyer (1987, 1990, 1991, 1992) solved the criterion problem by comparing subjects' self-evaluations of performance to their actual performance. The difference between self-evaluations and performance indicates the level of accuracy of self-perceptions. Beyer found that women's self-evaluations are not only lower than men's but are in fact inaccurately low on masculine gender-typed tasks.

Unlike research on gender, research on depression and self-esteem has recognized the importance of assessing the accuracy of self-perceptions. Some theories of depression (Beck, 1976) and self-esteem (Fitch, 1970) presumed that the self-perceptions of depressives and low self-esteem individuals were negatively distorted. However, many researchers have found evidence for depressive realism, also called the "sadder but wiser" phenomenon. For example, depressives are more accurate than nondepressives in their evaluations of social competence (Lewinsohn, Mischel, Chaplin, & Barton, 1980; McNamara & Hackett, 1986), recall of their toddlers' negative behaviors (Lovejoy, 1991), estimates of future success and failure (Alloy & Ahrens, 1987; Golin, Terrell, & Johnson, 1977; Golin, Terrell, Weitz, & Drost, 1979), estimates of positive and negative events that might happen to them (Crocker, Alloy, & Kayne, 1988), assessments of the degree of control over external stimuli (Abramson & Alloy, 1981; Alloy & Abramson, 1979, 1982; Alloy, Abramson, & Kossman, 1985; Alloy, Abramson, & Viscusi, 1981; Dobson & Franche, 1989; Glass, McKnight, & Valdimarsdottir, 1993; Martin, Abramson, & Alloy, 1984; Vazquez, 1987), and estimates of the extent of received punishment (Nelson & Craighead, 1977). Depressives also show less bias towards recalling flattering self-descriptions than nondepressives (Rude, Krantz, & Rosenhan, 1988). In addition, compared to nondepressives, mildly depressed individuals are more sensitive to changes in reward contingencies (Rosenfarb, Burker, Morris, & Cush, 1993).

However, DeMonbreun and Craighead (1977), Dobson and Shaw (1981), Gotlib (1981, 1983), and Wener and Rehm (1975) found no evidence for depressive realism. With the exception of Wener and Rehm (1975) these studies used psychiatric patients as subjects rather than dysphoric undergraduates. Thus, a boundary condition for depressive realism may be the severity of depression: Only mildly depressed individuals may show depressive realism with more depressed individuals showing negative biases.

Like mild depressives, low self-esteem individuals have more accurate self-evaluations of performance than do high self-esteem individuals (Shrauger & Kelly, 1988; Shrauger & Terbovic, 1976). For example, research on self-assessments of attractiveness indicates that most high self-esteem subjects, and males in general, tend to

overestimate their own attractiveness (Gurman & Balban, 1990). To summarize, research suggests that mildly depressed and low self-esteem individuals tend to have realistic self-perceptions. Studies on the effect of gender on the accuracy of self-perceptions indicate that on masculine tasks women tend to have overly negative self-perceptions. The present research will investigate the accuracy of depressives' self-evaluations of performance as a further test of the depressive realism hypothesis. In addition, gender differences in the accuracy of self-evaluations will be assessed.

Why might women inaccurately assess i.e., underestimate their performance on masculine tasks? There are reasons to believe that women and men fall prey to different self-perception biases. According to self-consistency theory "people interpret and judge their achievements and abilities in ways congruent with prior self-conceptions" (Jussim, Coleman, & Nassau, 1987, p. 95). Thus, a person's expectancy biases how performance on a task is interpreted. This is especially true when there is some ambiguity regarding the quality of performance, as in the absence of feedback (Felson, 1981; Wells & Sweeney, 1986). Therefore, self-consistency should result in inaccurate self-evaluations on those occasions when expectancies do not coincide with performance.

Because women have low expectancies for masculine tasks (e.g., Beyer, 1987, 1990, 1991, 1992; Janman, 1987), self-consistency theory predicts that they should evaluate their performance negatively. Conversely, men's high expectations should lead to high self-evaluations. Interestingly, there seems to exist a gender difference in the strength of self-consistency tendencies. Beyer (1990, 1991, 1992) found that women were more prone to self-consistency tendencies than men. A replication of this finding will be attempted.

Self-consistent tendencies may be revealed in yet another way. People selectively attend to information that is consistent with their self-views and show superior recall for such information (Swann & Read, 1981). For example, gender-consistent information is more likely to be recognized than gender-inconsistent information (Markus, Crane, Bernstein, & Sidali, 1982; Stangor, 1988). Patients suffering from panic attacks, show a bias toward threat stimuli compared to controls (Cloitre & Liebowitz, 1991). When processing self-referent material, depressives recall more negative and self-esteem threatening words or memories than positive words or memories, while normal controls recall more positive than negative or self-esteem threatening words (Bellew & Hill, 1990; Bradley & Mathews, 1983; Kuiper, Olinger, MacDonald, & Shaw, 1985; Lloyd & Lishman, 1975; McDowall, 1984). Thus, in many situations schema-consistent information seems to be attended to and recalled better than schema-inconsistent information.

It is possible that when evaluating their performance on masculine gender-typed tasks, women's recall of previously answered questions is biased by preexisting negative self-perceptions. Conceivably, if women have negative views or expectations to begin with, they may subsequently remember mostly those questions they believe they answered incorrectly, whereas men remember the questions they believe they answered correctly. This process could also bias women towards underestimation of their performance.

This experiment also was designed to determine whether women's inaccurately low self-evaluations would only be manifested when making overall self-evaluations or also at a more fundamental level, after assessing their performance for each individual question of a test. In other words, if subjects have to constantly monitor their performance, does this increase the accuracy of their self-evaluations?

In summary, this experiment tested the following hypotheses: 1. Depressed subjects are more accurate in their self-perceptions than nondepressed subjects (depressive realism). 2. Self-consistency tendencies can predict gender differences in self-evaluations. Women are hypothesized to show stronger self-consistency tendencies than men. 3. On the masculine task, women remember more of the questions they answered incorrectly than do men. 4. Subjects will be more accurate when constantly monitoring their performance than when only evaluating their overall performance. However, women's lower self-evaluations will be evident already when evaluating their performance on individual questions.

Method

Subjects. Subjects were 293 female and 174 male students at the University of Wisconsin-Parkside.

Tasks. Subjects were presented with either a feminine, masculine, or neutral gender-typed task, each containing 30 multiple-choice questions. The masculine task contained mathematics questions, the feminine task questions on grammar and syntax, and the neutral task questions on history and geography. Based on pretesting, the masculine and feminine tasks were constructed so that both genders would answer approximately 75% of the questions correctly on the gender-congruent task and 60% on the gender-incongruent task. The neutral task was constructed so that both genders would answer approximately 75% of the questions correctly.

Procedure. Subjects were randomly assigned to conditions and tasks. Subjects filled out the Beck Depression Inventory (BDI) first and subsequently were given information about the task they were about to perform (given sample questions, information about the number of questions, length of time available, etc.).

In the nonmonitoring condition subjects stated performance expectancies, performed the task, then estimated the number of correctly answered questions (self-evaluation) without receiving feedback regarding their performance. They then had to recall as many of the questions that had appeared on the task as possible and indicate for each recalled question, whether they believed they had answered that question correctly or incorrectly.

There was only one addition to this procedure in the monitoring condition. Immediately after answering each of the 30 multiple-choice questions, subjects stated how confident they were of having answered that particular question correctly. Confidence ratings could range from 0% to 100% sure. Only after completing these 30 ratings of question confidence did these subjects evaluate their overall performance. Thus, in the monitoring condition subjects were forced to constantly pay attention to their performance.

Results

Results were analyzed with 3 (gender-typedness of task: feminine, masculine, or neutral) x 2 (monitoring condition: monitoring vs. nonmonitoring) x 2 (gender) analyses of variance (ANOVAs).

Depressive realism. Accuracy scores were regressed on depression scores, gender, gender-typedness of task, monitoring condition, and all the interaction terms in multiple regression analyses. Both gender and monitoring condition interacted with depression scores, $F(1, 458) = 4.32$, $p < .04$, $F(1, 458) = 3.79$, $p < .06$, respectively. Therefore separate regression analyses for males and females and monitoring conditions were performed.

Subjects' depression scores significantly predicted accuracy in the monitoring condition of the feminine task, $F(1, 74) = 5.93$, $p < .02$, and marginally for the neutral nonmonitoring condition, $F(1, 58) = 3.60$, $p < .07$. The higher a subject's depression score, the more s/he underestimated performance. Subjects' depression scores significantly predicted their accuracy scores in the masculine monitoring condition, $F(1, 88) = 5.80$, $p < .02$. The higher a subject's depression score, the more accurately the subject evaluated her/his performance. In the feminine and masculine nonmonitoring conditions and the neutral monitoring condition depression scores did not significantly predict accuracy, $F(1, 83) < 1$, $F(1, 83) = 1.81$, $p < .19$, $F(1, 66) < 1$, respectively. Thus, with the exception of the masculine monitoring task, depression was not related to accurate self-evaluations, which represents a problem for the depressive realism hypothesis.

Accuracy of self-evaluations. Accuracy of self-evaluations was assessed by subtracting performance from self-evaluation scores. Positive discrepancies indicate overestimations, negative numbers underestimations, and scores around zero indicate accuracy in self-evaluations. 3 x 2 x 2 ANOVAs were conducted on accuracy scores. Because the effect of monitoring condition was not significant, $F(1, 460) < 1$, it was omitted from further analyses of accuracy scores. To determine whether an accuracy score is significantly different from zero, and whether there is a significant gender difference in accuracy for each task, a repeated measures ANOVA was computed with performance and self-evaluations as within-subjects factors and gender as between-subjects factor.

Women and men did not differ significantly in the accuracy of self-evaluations on the feminine task, $F(1, 159) < 1$. Both genders significantly underestimated their performance, $F(1, 159) = 26.53$, $p < .0001$. The gender difference in accuracy on the neutral task was marginally significant, $F(1, 126) = 3.35$, $p < .07$. Whereas men's self-evaluations were accurate, $F(1, 49) = 2.47$, $p < .13$, women significantly underestimated their performance, $F(1, 77) = 15.13$, $p < .0001$. Both males and females significantly overestimated their performance on the masculine task, $F(1, 67) = 45.36$, $p < .0001$, $F(1, 106) = 5.94$, $p < .02$, respectively. However, as predicted, the gender difference in accuracy of self-evaluations was significant on this task, $F(1, 173) = 5.00$, $p < .03$.

Self-consistency hypothesis. Self-evaluation was regressed on expectancy, performance, gender, gender-typedness of task, monitoring condition, and all the interaction terms in multiple regression analyses. Because the effect of monitoring condition was not significant, $F(1, 446) < 1$, it was omitted from further analyses. Gender interacted with expectancies to a significant degree, $F(1, 453) = 8.40$, $p < .004$, making separate regression analyses for males and females advisable.

Performance and expectancies accounted for significant amounts of variance in self-evaluations on the feminine task for women, $F(1, 102) = 33.67$, $p < .0001$, $F(1, 102) = 19.04$, $p < .0001$, respectively. These effects were present for men also, albeit to a considerably smaller extent: Performance significantly predicted self-evaluations, $F(1, 52) = 4.19$, $p < .05$, but expectancies were only marginally significant predictors of self-evaluations, $F(1, 52) = 3.77$, $p < .06$.

Performance and expectancies accounted for significant amounts of variance in self-evaluations on the neutral task for women, $F(1, 74) = 5.60$, $p < .03$, $F(1, 74) = 5.96$, $p < .02$, respectively. These effects were present for men also: Performance and expectancies significantly predicted self-evaluations, $F(1, 46) = 7.30$,

$p < .01$, $F(1, 46) = 5.27$, $p < .03$, respectively.

Performance and expectancies accounted for significant amounts of variance in self-evaluations on the masculine task for women, $F(1, 103) = 57.12$, $p < .0001$, $F(1, 103) = 48.57$, $p < .0001$, respectively. As hypothesized, for men performance significantly predicted self-evaluations, $F(1, 64) = 86.23$, $p < .0001$, whereas expectancy was not a significant predictor of self-evaluations, $F(1, 64) = 1.61$, $p < .21$. Thus, as hypothesized, women relied more on expectancies when evaluating their performance, i.e., were relatively more influenced by self-consistency tendencies, than men, especially on the masculine task.

Gender differences in recall. It was hypothesized that on the masculine task women would be more likely to recall questions they had answered incorrectly. This was tested by means of a 3 x 2 x 2 analysis of covariance with subjects' performance on a task as covariate and the number of recalled questions which had been answered incorrectly divided by total recall as the dependent variable. This analysis takes differences in total recall and actual performance into account. As recommended by Winer, Brown, and Michels (1991), arcsine transformations were performed on these data.

Because the effect of monitoring condition was not significant, $F(1, 436) < 1$, it was omitted from further analyses. On neither the feminine, nor the neutral task was the gender difference in the proportion of recalled questions which had been answered incorrectly significant, $F(1, 143) < 1$, $F(1, 125) < 1$, respectively. On the masculine task the gender difference was borderline significant, $F(1, 163) = 3.21$, $p < .08$. Thus, even with performance covaried out, women were somewhat more likely than men to recall questions which they had answered incorrectly.

A further analysis was conducted to analyze gender differences in recall in even more detail. Those recall scores representing errors in judgment were analyzed. Two such errors are possible: 1. thinking you answered a question incorrectly when in fact you answered it correctly (number of recalled questions that had been answered correctly and which the subject thought s/he had answered incorrectly divided by that subject's total recall of questions which had in fact been answered correctly) and 2. thinking you answered a question correctly when in fact you answered it incorrectly (number of recalled questions that had been answered incorrectly and which the subject thought s/he had answered correctly divided by that subject's total recall of questions which had in fact been answered incorrectly). Scores were corrected for differences in recall. Arcsine transformations were performed on these data. These data were analyzed with a 3 x 2 x 2 ANOVA.

Because of the presence of several interactions among the variables for the first recall variable (answered correctly but subject thought question was answered incorrectly divided by subject's total recall of questions which had been answered correctly), the data were analyzed separately by task and monitoring condition. In the feminine monitoring condition, the gender difference in recall was not significant, $F(1, 64) < 1$. In the feminine, neutral, and masculine nonmonitoring conditions the gender difference was not significant, $F(1, 80) = 2.48$, $p < .12$, $F(1, 58) = 2.13$, $p < .15$; $F(1, 78) < 1$, respectively. In the masculine and neutral monitoring conditions, some evidence for differential recall was found. The gender difference in recall was borderline significant for the neutral task, $F(1, 66) = 2.89$, $p < .10$, and significant for the masculine task, $F(1, 85) = 4.15$, $p < .05$. As predicted, women more often than men judged correctly answered questions as having been answered incorrectly.

Because monitoring condition was not a significant factor in the analyses of the second recall variable (questions which had been answered incorrectly, but for which the subject thought the question was answered correctly), it was dropped from further analyses of this dependent variable. The gender difference in recall was not significant for the feminine task, $F(1, 146) < 1$, the neutral task, $F(1, 126) = 1.70$, $p < .20$, or the masculine task, $F(1, 165) = 1.13$, $p < .29$.

Thus, on the masculine task information on failure was more mentally available to women than men. Interestingly, on the masculine and neutral tasks information on misperceived (as opposed to actual) failure was more mentally available for women than for men. This differential recall plus women's reliance on self-consistency may explain women's lower estimations of performance.

Gender differences in False Alarms and Misses. Subjects' confidence statements for each individual question in the monitoring condition were analyzed in terms of the proportion of false alarms and misses. A false alarm is an incorrectly answered question for which the subject was highly confident that it was answered correctly. A miss is a correctly answered question for which the subject showed little confidence. Misses were transformed into proportions by dividing each by a subject's performance score. False alarms were transformed into proportions by dividing each by a subject's number of incorrectly answered questions. An arcsine transformation was performed on these proportion data. A high proportion of false alarms indicates overly high confidence. A high proportion of misses indicates overly low confidence.

On the feminine task there was no gender difference in the proportions of false alarms and misses, $Fs(1, 1, 75 < 1$. On the neutral and masculine tasks, men had significantly more false alarms than women i.e., more

frequently had high confidence when they answered a question incorrectly, $F(1, 66) = 9.32$, $p < .003$, $F(1, 87) = 4.08$, $p < .05$, respectively. Women had significantly more misses on the masculine task than men i.e., more frequently had low confidence when they answered a question correctly, $F(1, 89) = 10.75$, $p < .001$.

Discussion

This study found little evidence to support the depressive realism hypothesis. In only 3 out of 6 conditions was depression significantly related to accuracy of self-perception. However, contrary to the depressive realism hypothesis, in 2 out of those 3 conditions depression was related to less accurate self-evaluations.

The hypothesis that gender differences in accuracy depend on the gender-typedness of the task received support. The only task for which the gender difference in accuracy reached significance was the masculine task. The results for the neutral task were of borderline significance. However, a surprising quirk in the data was found. It was hypothesized that women underestimate their performance on the masculine task. In fact they overestimated their performance albeit to a significantly smaller extent than men overestimated their performance. This result represents a departure from the results of four previous experiments in which women underestimated their performance on the masculine task (Beyer, 1987, 1990, 1991, 1992). With 20/20 hindsight this result can be explained by the nature of the task that was employed in this experiment. Whereas the previous four experiments employed sports and politics questions, the present experiment used math questions for the masculine task. The problem with math questions is that in an indirect way feedback on the accuracy of one's solution to a problem is provided: after working through a problem, the answer either matches or does not match one of the provided response alternatives. When there is a match, the subject is likely to assume that s/he answered the question correctly. However, by design, several of the response alternatives were meant to mislead subjects by matching incorrect solutions. Thus, a subject might produce a match with an incorrect response alternative and assume that this match meant that the question was answered correctly. This should lead to an inaccurately high self-evaluation. No such indirect feedback on performance is provided by sports and politics questions nor by the feminine (English language) and neutral (history and geography) questions of this experiment. Thus, math questions represent a less than ideal kind of masculine items. However, one should not lose sight of the fact that consistent with prior experiments, the only significant gender difference in the accuracy of self-evaluations was found on the masculine task.

Expectancies had a significant effect on self-evaluations, demonstrating the existence of self-consistency tendencies. As predicted, performance was the best predictor of self-evaluations on all tasks. For women unlike for men, expectancies played almost as important a role as did performance in predicting self-evaluations. On the masculine task, men's self-evaluations were predicted only by performance, not by expectancies. This indicates that men's self-evaluations are more guided by their performance and women's self-evaluations are guided to a considerable extent by self-consistency. Because women's initial expectancies tend to be lower than men's, especially on masculine tasks, this reliance on self-consistency when evaluating performance results in lower self-evaluations.

As predicted, gender differences in recall of incorrectly answered questions were found on the masculine task. Thus, on the masculine task more information on failure was mentally available to women than men. What is even more interesting, however, is that on the masculine and neutral tasks information on misperceived (as opposed to actual) failure was more mentally available for women than for men. Thus, women's recall compared to men's was biased in a negative direction. It is important to note that this gender difference in biased recall was not evident on the feminine task. This biased recall plus women's reliance on self-consistency may explain women's lower estimations of performance.

Contrary to the hypothesis, when subjects had to state their confidence for each individual question, they were no more accurate in their self-evaluations than subjects who were not forced to constantly monitor their performance. This points out that simply paying more attention to one's performance does not enhance accuracy.

As hypothesized, when examining the data regarding the presence of biased self-evaluations at the level of individual questions data consistent with the above results were found. On the feminine task for which no gender difference in the accuracy of overall self-evaluations was found, no gender difference in the accuracy of self-evaluations at the level of individual questions was found. However, on the neutral and masculine tasks for which gender differences in overall self-evaluations were present, gender differences in self-evaluations at the level of individual questions were found: Men more frequently had high confidence when they answered a question incorrectly than did women, whereas women more frequently had low confidence when they answered a question correctly than did men. Thus, on the masculine and neutral tasks, biased self-perceptions were already operating when evaluating performance on each individual question. This seems to indicate that

women's self-perceptions are not only biased when mentally averaging performance on individual questions into one overall self-evaluation, but are biased already at the more fundamental level of evaluating performance on individual questions of a task. This statement needs qualification, however. This occurs only for masculine and neutral tasks.

Why is it important to demonstrate that women's self-evaluations are frequently inaccurately low? Positive self-perceptions, even if they are inaccurately high, are related to psychological health (Snyder, 1989; Taylor, Collins, Skokan, & Aspinwall, 1989), improved motivation, and task persistence (Abramson & Alloy, 1981). Low perceptions of performance negatively affect performance (Elliott & Dweck, 1988), persistence (Elliott & Dweck, 1988), expectancies for future performance (Phillips, 1987), aspirations (Phillips, 1984), and affect (Elliott & Dweck, 1988). Negative self-evaluations combined with stressors can lead to depression (Brown, Andrews, Bifulco, & Veiel, 1990; Brown, Andrews, Harris, Adler, & Bridge, 1986; Brown, Bifulco, & Andrews, 1990).

Thus, women's inaccurately low self-evaluations may have damaging consequences. For example, females who received high grades in math courses, but nevertheless had low expectancies for future grades, did not enroll in advanced math courses (Lantz & Smith, 1981). For males, it was only poor performance which led to an avoidance of math courses. Thus, men's future math taking behavior could be predicted by grades (performance), whereas women's math taking behavior could be predicted by low expectancies (self-consistency). This study of naturalistic behavior nicely supports the findings of this experiment regarding differential emphasis on self-consistency and performance by men and women.

But why do females who received superior grades in math develop low expectancies for future math grades? The present experiment suggests that one reason may be females' reliance on self-consistency. If females have low expectancies for math performance to begin with, they are likely to inaccurately assess their performance in math. If inaccurately low self-evaluations affect future expectancies negatively, females are unlikely to take more math in the future. This may partially account for the underrepresentation of women in math (Eccles, 1987). Still, some vexing questions remain. Why would females have low expectancies for math to begin with and why does objective feedback such as grades not alter females' expectancies for future math courses?

Females are socialized to be modest, whereas males are taught to be confident regarding academic achievements (Phillips, 1987). Many parents have inaccurately low perceptions of their daughters' ability in such areas as math. These low perceptions eventually come to be shared by their daughters (Parsons, Adler, & Kaczala, 1982). Thus, females learn from parents and society to underestimate their competence. As this research has demonstrated, females learn their lessons well, i.e., women indeed tend to have lower expectancies than men. Unfortunately, because of females' reliance on self-consistency, once they have learned their lesson (to have low expectancies), they have difficulty unlearning it.

It is not too difficult to believe that when feedback about actual performance is absent, such as in the present research, biases such as self-consistency could come into play. But what about those cases where clear, unambiguous feedback regarding performance is available such as in the above-mentioned study by Lantz and Smith (1982). Why do so many females who receive feedback regarding performance in the form of high grades in math believe that they will do poorly in the future? The recall data may provide some insight here. On the masculine task, women were more likely than men to recall questions they believed they answered incorrectly. Such biased recall of negative information is likely to affect self-evaluations. If a relatively high proportion of information on believed failure is mentally available when evaluating one's performance, this should negatively bias self-evaluations. Many of us have known individuals who, after receiving feedback on their performance, focus on and remember the tiny bit of criticism rather than the overwhelming amount of praise. Perhaps females who receive high grades in math focus on the negative aspects of their performance (mistakes) rather than the positive aspects (high grades), therefore perceive their performance as failure and avoid math in the future.

Because of the serious implications of underestimations of performance for self-confidence and psychological health, more attention should be devoted to the investigation of gender differences in the accuracy of self-evaluations. Such research will not only elucidate the underlying processes of self-evaluation biases and therefore be of theoretical interest, but will also be of practical value by suggesting ways of reducing women's underestimations of performance.

References

Abramson, L. Y., & Alloy, L. B. (1981). Depression, nondepression, and cognitive illusions: Reply to Schwartz. Journal of Experimental Psychology: General, 110, 436-447.

Alloy, L. B., & Abramson, L. Y. (1979). Judgment of contingency in depressed and nondepressed students: Sadder but wiser? Journal of Experimental Psychology: General, 108, 441-485.

Alloy, L. B., & Abramson, L. Y. (1982). Learned helplessness, depression, and the illusion of control. Journal of Personality and Social Psychology, 42, 1114-1126.

Alloy, L. B., Abramson, L. Y., & Kossman, D. A. (1985). The judgment of predictability in depressed and nondepressed college students. In F. R. Brush & J. B. Overmier (Eds.), Affect, conditioning and cognition: Essays on the determinants of behavior (pp. 229-246). Hillsdale, N.J.: Erlbaum.

Alloy, L. B., Abramson, L. Y., & Viscusi, D. (1981). Induced mood and the illusion of control. Journal of Personality and Social Psychology, 41, 1129-1140.

Alloy, L. B., & Ahrens, A. H. (1987). Depression and pessimism for the future: Biased use of statistically relevant information in predictions for self versus others. Journal of Personality and Social Psychology, 52, 366-378.

Beck, A. T. (1976). Cognitive therapy and the emotional disorders. New York: International Universities Press.

Bellew, M., & Hill, A. B. (1990). Negative recall bias as a predictor of susceptibility to induced depressive mood. Personality and Individual Differences, 11, 471-480.

Beyer, S. (1987). Expectancies of success: Are women more accurate than men? Unpublished Master's Thesis, University of Oregon.

Beyer, S. (1990). Gender differences in the accuracy of self-evaluations of performance. Journal of Personality and Social Psychology, 59, 960-970.

Beyer, S. (1991). The effect of self-consistency on gender differences in the accuracy of self-evaluations of performance. Dissertation, University of Oregon.

Beyer, S. (1992, August). Self-consistency and gender differences in the accuracy of self-evaluations. Paper presented at the annual convention of the American Psychological Association. Washington, D.C.

Bradley, B., & Mathews, A. (1983). Negative self-schemata in clinical depression. British Journal of Clinical Psychology, 22, 173-181.

Brown, G. W., Andrews, B., Bifulco, A., & Veiel, H. (1990). Self-esteem and depression: I. Measurement issues and prediction of onset. Social Psychiatry and Psychiatric Epidemiology, 25, 200-209.

Brown, G. W., Andrews, B., Harris, T., Adler, Z., & Bridge, L. (1986). Social support, self-esteem and depression. Psychological Medicine, 5, 813-831.

Brown, G. W., Bifulco, A., & Andrews, B. (1990). Self-esteem and depression: III. Aetiological issues. Social Psychiatry and Psychiatric Epidemiology, 25, 235-243.

Carr, P. G., Thomas, V. G., & Mednick, M. T. (1985). Evaluation of sex-typed tasks by black men and women. Sex Roles, 13, 311-316.

Cloitre, M., & Liebowitz, M. R. (1991). Memory bias in panic disorder: An investigation of the cognitive avoidance hypothesis. Cognitive Therapy and Research, 15, 371-386.

Crocker, J., Alloy, L. B., & Kayne, N. T. (1988). Attributional style, depression, and perceptions of consensus for events. Journal of Personality and Social Psychology, 54, 840-846.

Daubman, K. A., Heatherington, L., & Ahn, A. (1992). Gender and the self-presentation of academic achievement. Sex Roles, 27, 187-204.

DeMonbreun, B. G., & Craighead, W. E. (1977). Distortion of perception and recall of positive and neutral feedback in depression. Cognitive Therapy and Research, 1, 311-329.

Dobson, K., & Franche, R.-L. (1989). A conceptual and empirical review of the depressive realism hypothesis. Canadian Journal of Behavioral Sciences, 21, 419-433.

Dobson, K. S., & Shaw, B. F. (1981). The effects of self-correction on cognitive distortions in depression. Cognitive Therapy and Research, 5, 391-403.

Eccles, J. S. (1987). Gender roles and women's achievement-related decisions. Psychology of Women Quarterly, 11, 135-172.

Elliott, E. S., & Dweck, C. S. (1988). Goals: An approach to motivation and achievement. Journal of Personality and Social Psychology, 54, 5-12.

Elliot, A. J., & Harackiewicz, J. M. (1994). Goal setting, achievement orientation, and intrinsic motivation: A mediational analysis. Journal of Personality and Social Psychology, 66, 968-980.

Erkut, S. (1983). Exploring sex differences in expectancy, attribution, and academic achievement. Sex Roles, 9, 217-231.

Felson, R. B. (1981). Ambiguity and bias in the self-concept. Social Psychology Quarterly, 44, 64-69.
Fitch, G. (1970). Effects of self-esteem, perceived performance, and choice on causal attributions. Journal of Personality and Social Psychology, 16, 311-315.
Glass, D. C., McKnight, D., & Valdimarsdottir, H. (1993). Depression, burnout, and perceptions of control in hospital nurses. Journal of Consulting and Clinical Psychology, 61, 147-155.
Golin, S., Terrell, F., & Johnson, B. (1977). Depression and the illusion of control. Journal of Abnormal Psychology, 86, 440-442.
Golin, S., Terrell, F., Weitz, J., & Drost, P. L. (1979). The illusion of control among depressed patients. Journal of Abnormal Psychology, 88, 454-457.
Gotlib, I. H. (1981). Self-reinforcement and recall: Differential deficits in depressed and nondepressed psychiatric inpatients. Journal of Abnormal Psychology, 90, 521-530.
Gotlib, I. H. (1983). Perception and recall of interpersonal feedback: Negative bias in depression. Cognitive Therapy and Research, 7, 399-412.
Gurman, E. B., & Balban, M. (1990). Self-evaluations of physical attractiveness as a function of self-esteem and defensiveness. Journal of Social Behavior and Personality, 5, 575-580.
Janman, K. (1987). Achievement motivation theory and occupational choice. European Journal of Social Psychology, 17, 327-346.
John, O. P., & Robins, R. W. (1994). Accuracy and bias in self-perception: Individual differences in self-enhancement and the role of narcissism. Journal of Personality and Social Psychology, 66, 206-219.
Jussim, L., Coleman, L., & Nassau, S. (1987). The influence of self-esteem on perceptions of performance and feedback. Social Psychology Quarterly, 50, 95-99.
Kuiper, N. A., Olinger, L. J., MacDonald, M. R., & Shaw, B. F. (1985). Self-schema processing of depressed and nondepressed content: The effects of vulnerability to depression. Social Cognition, 3, 77-93.
LaNoue, J. B., & Curtis, R. C. (1985). Improving women's performance in mixed-sex situations by effort attributions. Psychology of Women Quarterly, 9, 337-356.
Lantz, A. E., & Smith, G. P. (1981). Factors influencing the choice of nonrequired mathematics courses. Journal of Educational Psychology, 73, 825-837.
Lewinsohn, P. M., Mischel, W., Chaplin, W. C., & Barton, R. (1980). Social competence and depression: The role of illusory self-perceptions. Journal of Abnormal Psychology, 89, 203-212.
Lloyd, G. G., & Lishman, W. A. (1975). Effect of depression on the speed of recall of pleasant and unpleasant experiences. Psychological Medicine, 5, 173-180.
Lovejoy, M. C. (1991). Maternal depression: Effects on social cognition and behavior in parent-child interactions. Journal of Abnormal Child Psychology, 19, 693-706.
Martin, D. J., Abramson, L. Y., & Alloy, L. B. (1984). Illusion of control for self and others in depressed and nondepressed college students. Journal of Personality and Social Psychology, 46, 125-136.
McDowall, J. (1984). Recall of pleasant and unpleasant words in depressed subjects. Journal of Abnormal Psychology, 93, 401-407.
McNamara, K., & Hackett, G. (1986). Gender, sex-type and cognitive distortion: Self-perceptions of social competence among mild depressives. Social Behavior and Personality, 14, 113-121.
Meehan, A. M., & Overton, W. F. (1986). Gender differences in expectancies for success and performance on Piagetian spatial tasks. Merrill-Palmer Quarterly, 32, 427-441.
Mura, R. (1987). Sex-related differences in expectations of success in undergraduate mathematics. Journal for Research in Mathematics Education, 18, 15-24.
Nelson, R. E., & Craighead, W. E. (1977). Selective recall of positive and negative feedback, self-control behaviors, and depression. Journal of Abnormal Psychology, 86, 379-388.
Parsons, J. E., Adler, T. F., & Kaczala, C. M. (1982). Socialization of achievement attitudes and beliefs: Parental influences. Child Development, 53, 310-321.
Parsons, J. E., Meece, J. L., Adler, T. F., & Kaczala, C. M. (1982). Sex differences in attributions and learned helplessness. Sex Roles, 8, 421-432.
Phillips, D. (1984). The illusion of incompetence among academically competent children. Child Development, 55, 2000-2016.
Phillips, D. (1987). Socialization of perceived academic competence among highly competent children. Child Development, 58, 1308-1320.
Rosenfarb, I. S., Burker, E. J., Morris, S. A., & Cush, D. T. (1993). Effects of changing contingencies on the behavior of depressed and nondepressed individuals. Journal of Abnormal Psychology, 102, 642-646.
Rude, S. S., Krantz, S. E., & Rosenhan, D. L. (1988). Distinguishing the dimensions of valence and belief

consistency in depressive and nondepressive information processing. Cognitive Therapy and Research, 12, 391-407.

Rustemeyer, R. (1982). Wahrnehmungen eigener Faehigkeiten bei Jungen und Maedchen. Frankfurt: Peter Lang.

Shrauger, J. S., & Kelly, R. J. (1988). Global self-evaluation and changes in self-description as a function of information discrepancy and favorability. Journal of Personality, 56, 709-728.

Shrauger, J. S., & Terbovic, M. L. (1976). Self-evaluation and assessments of performance by self and others. Journal of Personality and Social Psychology, 44, 564-572.

Snyder, C. R. (1989). Reality negotiation: From excuses to hope and beyond. Journal of Social and Clinical Psychology, 8, 130-157.

Swann, W. B., Jr., & Reid, S. J. (1981). Acquiring self-knowledge: The search for feedback that fits. Journal of Personality and Social Psychology, 41, 1119-1128.

Taylor, S. E., Collins, R. L., Skokan, L. A., & Aspinwall, L. G. (1989). Maintaining positive illusions in the face of negative information: Getting the facts without letting them get to you. Journal of Social and Clinical Psychology, 8, 14-129.

Vazquez, C. (1987). Judgment of contingency: Cognitive biases in depressed and nondepressed subjects. Journal of Personality and Social Psychology, 52, 419-431.

Wells, L. E., & Sweeney, P. D. (1986). A test of three models of bias in self-assessment. Social Psychology Quarterly, 49, 1-10.

Wener, A. E., & Rehm, L. P. (1975). Depressive affect: A test of behavioral hypotheses. Journal of Abnormal Psychology, 84, 221-227.

Winer, B. J., Brown, D. R., & Michels, K. M. (1991). Statistical principles in experimental design (3rd ed.). New York: McGraw-Hill.